Queen's Economics Department Working Paper No. 74

AUTOCORRELATION IN EMPIRICAL STUDIES OF WAGE DETERMINATION

J.C.R. Rowley
Queen's University

D.A. Wilton
Queen's University

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

3-1972

AUTOCORRELATION IN EMPIRICAL STUDIES OF WAGE DETERMINATION

by

J. C. R. Rowley and D. A. Wilton

Discussion Paper No. 74

March 1972

The need for careful specification in econometric analyses is well-known. If some component of a particular specification is incorrect then conventional inferential procedures may be invalid. In particular, the presence of autocorrelation in linear economic models may lead to the use of inappropriate formulae[1] for estimates of standard errors of estimated coefficients and for F-statistics which provide the bases for tests of significance. Concern over this problem has generally been represented in recent years by two checks of sensitivity; namely, the calculation of Durbin-Watson d-statistics and the use of either autoregressive transformations of the type introduced by Cochrane and Orcutt [1949] or similar approaches such as the Hildreth-Lu [1960] scan procedure which can be associated with maximum likelihood methods.

In many instances, these checks must be characterized as "tokenisms" which prove to be inadequate when more attention is given to the underlying structures, for which the models are intended to provide simple representations. Econometric studies of wage-determination provide an excellent illustration of certain specific deficiencies which may be introduced by conventional checks of sensitivity. Clarification of the quantitative significance of these deficiencies depends critically upon adequate investigation of institutional features of the labour market and we cannot claim to have achieved this final goal. However, incompleteness of knowledge does not preclude a demonstration that current practices in empirical studies of wage-determination leave much to be desired. Two well-known studies of Phillips' curves, those of Perry [1966, 1970] and Bodkin et al. [1966], provide the bases for discussion. These were selected on the grounds that they have enjoyed immense influence and appear to be well-conducted in many other respects. The following quotation from the Canadian study of Bodkin et al. serves as a paradigm for innumerable other comments of a similar type.

> "Another difficulty, which was encountered in the wage adjustment equations, is the problem of autocorrelated residuals....Although this phenomenon does not, in the usual case, lead to biased parameter estimates, it does usually entail an understatement of the standard errors (computed according to the usual formulas) of the sample regression coefficients; consequently, the level of statistical significance of the explanatory variables will, in the usual case, be overstated. Fortunately, this difficulty can be handled in a number of ways....we form autoregressive transformations for all of the variables and then run the wage adjustment regressions with the transformed variables. In general, this transformation of the variables succeeds in eliminating the observed autocorrelation in the sample residuals, as judged by the Durbin-Watson statistic."[2]

In fact, for a simple model which is consistent with the approaches of both Perry and Bodkin et al., it can be shown that the autoregressive transformation leads to estimators which are generally less efficient than certain other calculable estimators. Empirical results which are tabulated below suggest that, in particular cases, the former estimators are individually less efficient than those obtained by applying the least-squares method to the untransformed model. This criticism is independent of the comment by Kadiyala [1968] on the loss of efficiency due to the "elimination" of a single observation when the transformation is used. In addition, with one specification of Bodkin et al., the autoregressive transformation introduces asymptotic bias for estimators whereas this bias is wholly absent for estimators which are based on the untransformed model. This final assertion may be associated with a direct contradiction of a presumption due to Bodkin et al., and it invalidates use of some of their influential empirical results. After reference to the Durbin-Watson tabulations and to the alternative Theil-Nagar [1961] approach, they make the following incorrect statement.

> "Ordinarily, the presence of a lagged dependent variable vitiates such tests, but since the lagged dependent variable is dated four quarters previously to the current quarter, presumably the reasoning underlying these tests will continue to go through, provided one can assume that the autoregressive structure for the disturbances do not extend backward for more than three quarters."[3]

Most empirical studies of wage-determination, including those cited above, have either the absolute or relative annual change in an aggregate wage-index for their dependent variable and simple four-quarters moving-averages for explanatory variables. These particular specifications are peculiar to wage-studies and can only be explained by reference to the characteristics of the market for labour, especially the presence of collective bargaining and long-term agreements between labour and management. In the simplest description, all labour is covered by annual agreements and members of the labour force can be divided into four distinct micro groups according to the quarters in which they negotiate and receive annual revisions in their wage-levels. Since data for individual groups are seldom available in appropriate detail, "macro-data" are used and these may be recognized as moving-averages of the quantities with which theoretical discussants are concerned.[4] There are substantial expositionary gains if we use the matrix notation, made popular by Johnston [1971] and Goldberger [1964], to represent the micro-model and its macro-analogue. The former model may be represented by

$$(1) \qquad y = X\beta + u$$

where y and X contain observations for the micro-wage-change variable and untransformed explanatory variables respectively. The matrix X is assumed to be non-stochastic[5] and the elements of the vector of errors are assumed to form a stationary white-noise sequence; that is, they have zero means, constant variances σ^2 and are free from autocorrelation. Their dispersion matrix is, therefore, $\sigma^2 I$ where I is the identity matrix of appropriate order. The macro-model, which is the basis for empirical investigation, is represented by

$$(2) \qquad Gy = GX\beta + Gu$$

where G is a rectangular matrix with fixed elements such that the available measurements are Gy and GX. Elements of the vector Gy are observations for the change in an aggregative wage-index whereas those of GX are moving-averages of explanatory variables. Notice that if Gy and GX are known, then the elements of G must be known. Perry chose to define G as a matrix of order (n-3) by n for some n with unit elements for elements indexed by $\{(i, i+s)$ for $s = 0, 1, 2, 3\}$ and zero otherwise. This assumption of fourth-order moving averages is maintained throughout the discussion provided below. However, all of this discussion can be re-stated in terms of a general class[6] of choices for G. Perry's assumption is consistent with the approach of Bodkin et al. even though his explicit comments are not replicated in the Canadian study.

Suppose u_{0t} is a typical element of Gu, the vector of macro-errors. Then, these errors for the transformed model (2) have zero means and their autocovariance function is the sequence of expected products $\{E(u_{0t}\, u_{0t+s})$ for $s = 0, 1, 2, \ldots\}$ as a function of the lag between errors, s. Let these products be denoted by $\{\gamma_s\}$. For all points of observation, Perry's assumption implies that

$$(3) \qquad \gamma_s = \gamma_{-s} = (4 - s)\sigma^2 \qquad \text{for } s = 0, 1, 2, 3$$

and zero otherwise. The Yule-Slutsky effect of the aggregative transformation for the errors is represented by the positivity of γ_1, γ_2 and γ_3. If the values of the autocovariance function for the transformed errors are tabulated in a dispersion matrix $\sigma^2 GG'$, this is a Laurent matrix with a band of positive elements about the principal diagonal and zero elements elsewhere. The non-zero elements are known up to the scalar constant σ^2 as a consequence of Perry's assumption for the weights of the moving-average matrix G. With these specifications, the aggregative model for wage-determination satisfies the conditions of the generalized classical linear statistical model for which Aitken's procedure provides the best, linear[7] and unbiased estimator of the parameters in the vector β.

The dispersion matrix $\sigma^2 GG'$ is markedly different from that associated with the original errors $\sigma^2 I$. This feature of the model is well known. However, Perry is incorrect when he asserts

> Even if the micro-errors satisfy the usual least-squares assumptions of being normally distributed with zero mean, finite variance, and uncorrelated with the explanatory variables, the macro-errors would not satisfy these conditions. This would not introduce asymptotic bias with least-squares estimates on macro-equations but it would impair the efficiency of the estimates and produce uncertain small-sample properties for the estimates.[8]

In fact, these errors are normally distributed with zero mean, finite variance and uncorrelated with the explanatory variables. Although the least-squares estimators are free from asymptotic bias and they are inefficient relative to Aitken's GLS estimators, their small-sample properties can be explicitly established.

If the explanatory variables for the transformed model include a lagged dependent variable so that GX has stochastic elements, these properties must be modified but the least-squares estimators may remain free from asymptotic bias if the lag for this variable exceeds three quarters. Perry's approach does not involve any lagged dependent variables but Bodkin et al. do include a four-quarters lagged dependent variable. This variable is contemporaneously uncorrelated with the macro-error and the method of least-squares may be appropriate. Clearly this result may have influenced the thinking of Bodkin et al. when they made the presumption cited in the second quotation.[9] Even if no lagged dependent variable is present, they are incorrect.

Since the nature of the labour market is known to result in dispersion matrices of the form $\sigma^2 GG'$ which necessarily indicate the presence of autocorrelation due to aggregation, there would appear to be no reason to check for its existence. Further, Durbin and Watson are quite explicit in their derivation of the bounds' test when they indicate how it should be used. The particular form of the d-statistic was chosen so that it would have substantial power in a simple comparison between uncorrelated errors and those generated by a first-order autoregressive processes. It was never intended to confirm the presence of moving-averages and tabulated values of the bounds for the statistic are of doubtful relevance in this context. Thus it is difficult to accept statements of the type:

> Finally the presence of first-order serial correlation in the residuals is evident in this fully specified form of this wage model. The Durbin-Watson statistic...is 1.2. As explained earlier, this is to be expected from the overlap in the independent variables that results from explaining wage changes over a one-year interval.[10]

The dispersion matrix $\sigma^2 GG'$ is, also, markedly different from those matrices associated with errors which are generated by autoregressive processes of any finite order. Thus the use of either the Cochrane-Orcutt transformation or the Hildreth-Lu scan procedure is questionable. In their study of British data, Dicks-Mireaux and Dow [1959] used the Cochrane-Orcutt transformation with 0.75 as the value prescribed for the "autoregressive parameter". This value was selected as the ratio of γ_1 to γ_0, the first "autocorrelation" of the macro-errors when Perry's assumption of unit values is made. Bodkin et al. rejected this value on the grounds that it "apparently 'overcorrected' in the sense of inducing negative first-order autocorrelation in the new error terms"[11]. (Their evidence for this rejection was the Durbin-Watson statistic!) They preferred the value 0.375. Although Perry [1966] only mentions the consequences of aggregation and applies the method of least-squares to the macro-model, his later study [1970] involves the scan procedure with 0.65 and 0.70 as alternative parameters for the implicit autoregressive transformation. In all these cases, the problem of autocorrelated errors is not eliminated. The principal effect of autoregressive transformations is the lengthening of the moving average for the errors. Let $\{u_{\rho t}\}$ represent the errors which result from the use of an autoregressive transformation with parameter ρ and let $\{u_{0t}\}$ represent the initial macro-errors, the elements of Gu.

$$(4) \qquad u_{0t} \equiv u_t + u_{t-1} + u_{t-2} + u_{t-3}$$

$$(5) \qquad u_{\rho t} = u_t + (1-\rho)u_{t-1} + (1-\rho)u_{t-2} + (1-\rho)u_{t-3} - \rho u_{t-4}$$

Dispersion matrices for these new errors will be "banded" Laurent matrices for all choices of ρ. For choices constrained within the closed interval [0,1], the autoregressive transformations yield autocovariances which satisfy the following inequalities.

$$(6) \qquad 0 < \gamma_{\rho s}/\gamma_{\rho 0} < \gamma_{0s}/\gamma_{00} \qquad \text{for } s = 1, 2, 3.$$

The autocorrelations $\{(\gamma_{\rho s}/\gamma_{\rho 0})$ for s = 1, 2, 3$\}$ are monotonically-decreasing positive functions of ρ within the interval and are zero for unit ρ (as indicated by Champernowne in his discussion of the paper by Dicks-Mireaux and Dow) but the autocovariance $\gamma_{\rho 4}$ is non-zero for all non-zero values of ρ.

(7) $\gamma_{\rho 0} / \sigma^2 = 4(1-\rho)^2 + 2\rho$

(8) $\gamma_{\rho s} / \sigma^2 = (4-s)(1-\rho)^2 \qquad \text{for } s = 1, 2, 3$

(9) $\gamma_{\rho 4} / \sigma^2 = -\rho$

For example, the non-zero autocorrelations for ρ equal to 0.375 and to 0.75 are {0.507, 0.338, 0.169, - 0.162} and {0.107, 0.071, 0.36, - 0.429} respectively. These formulae must be used if the efficiencies of the alternative estimates for β are to be calculated. Notice that the autoregressive transformations reduce three of the positive autocorrelations but, unfortunately, they introduce substantial negative fourth autocorrelations.

Before attention is restricted to comparisons of the efficiencies of different estimators, one other important consequence of autoregressive transformation must be indicated. We have already pointed out that the four-quarters lagged dependent variable in the specification of Bodkin et al. may not lead to asymptotic biases in the least-squares estimators for their model. Unfortunately, this lengthening of the moving-average for the errors will lead to asymptotic biases for their estimators whenever their data are subjected to either of the two popular approaches for autoregressive transformations. Contemporaneous correlation is introduced between the errors and the lagged dependent variable.

The Efficiencies of Alternative Estimators: Some Numerical Estimates[12]

The models of Perry [1970] and Bodkin et al. provide the bases for a number of empirical results which are tabulated below. Three alternative estimators are considered for each model. Least-squares estimators and Aitken's GLS estimators are denoted by b_0 and b_g respectively.

$$(10) \qquad b_0 \equiv A_0 Gy \qquad \text{where} \quad A_0 \equiv (X'G'GX)^{-1}X'G'$$

$$(11) \qquad b_g \equiv A_g Gy \qquad \text{where} \quad A_g \equiv \{X'G'(GG')^{-1}GX\}^{-1}X'G'(GG')^{-1}$$

Suppose y_0 is a vector of length (n-4) formed from the elements of Gy except for the first observation. Similarly, let y_{-1} represent the vector formed from the elements of Gy except for the final observation, and let X_0 and X_{-1} represent matrices formed from GX by delection of its first row and final row respectively. Then, we can define the data-matrix, $(X_\rho : y_\rho)$, used in obtaining estimators b_ρ after autoregressive transformations in terms of these vectors and matrices.

$$(12) \qquad X_\rho \equiv X_0 - \rho X_{-1}$$

$$(13) \qquad y_\rho \equiv y_0 - \rho y_{-1} \quad = \quad MGy$$

where M is defined implicitly by (13). Then, for any given value of ρ, b_ρ is given by

$$(14) \qquad b_\rho \equiv A_\rho Gy \qquad \text{where} \quad A_\rho \equiv (X_\rho' X_\rho)^{-1}X_\rho' M.$$

If the explanatory variables in the models of wage-determination are nonstochastic, all three estimators are unbiased with respect to β and we need some other criterion to distinguish between them. Computational simplicity would indicate b_0 but neither of the other estimators involves considerable computational difficulties for the sample sizes which are available. Efficiency is the secondary criterion which is most frequently invoked and this criterion is assumed to be of paramount importance for the remainder of this section.

The dispersion matrix, or variance-covariance matrix, of any estimator is represented by the symbol $\mathcal{D}(\cdot)$. Then, if σ^2GG' is the dispersion matrix of Gy,

$$(15) \qquad \mathcal{D}(b_0) = \sigma^2 A_0 GG'A_0'$$

$$(16) \qquad \mathcal{D}(b_g) = \sigma^2 A_g GG'A_g' \qquad \text{and}$$

$$(17) \qquad \mathcal{D}(b_\rho) = \sigma^2 A_\rho GG'A_\rho' \ .$$

Let $\Lambda(\cdot)$ represent the matrix obtained from any square matrix $(\cdot)$ by setting all of its off-diagonal elements equal to zero. Then, by the Gauss-Markov theorem,

$$(18) \qquad \Lambda(A_g GG'A_g') \leqq \Lambda(A_0 GG'A_0')$$

$$(19) \qquad \Lambda(A_g GG'A_g') \leqq \Lambda(A_\rho GG'A_\rho) \qquad \text{for all } \rho.$$

Further, a corollary of the theorem permits the following inequalities to be established for the "generalized variances" of collections of estimators, the determinants of the dispersion matrices for these estimators.

$$(20) \qquad \det\,(A_g GG'A_g') \leqq \det\,(A_0 GG'A_0')$$

$$(21) \qquad \det\,(A_g GG'A_g') \leqq \det\,(A_\rho GG'A_\rho').$$

Two concepts of "efficiency" are based upon $\Lambda(\cdot)$ and $\det(\cdot)$ for different dispersion matrices of estimators. These concepts are denoted $EF(I)_i$ and

$EF(II)_i$ for subscript i representing either the least-squares estimator (i is 0) or the autoregressive-transformation estimator, henceforth referred to as the Scan estimator for brevity (i is ρ).

$$(22) \quad EF(I)_0 \equiv \text{vec}\, \{\Lambda(A_0 GG'A_0')\, \Lambda^{-1}\, (A_g\, GG'A_g')\} ,$$

the column vector formed from the principal diagonal of the matrix $\{\cdot\}$.

$$(23) \quad EF(I)_\rho \equiv \text{vec}\, \{\Lambda(A_\rho GG'A_\rho')\, \Lambda^{-1}\, (A_g GG'A_g')\}$$

$$(24) \quad EF(II)_0 \equiv \det\, (A_0 GG'A_0') \,/\, \det\, (A_g GG'A_g')$$

$$(25) \quad EF(II)_\rho \equiv \det\, (A_\rho GG'A_\rho') \,/\, \det\, (A_g GG'A_g').$$

The first concept of efficiency concerns pair-wise comparisons of estimators for the same parameters. The ideal value of EF(I) would be a column vector of unit elements. The efficiencies of either the least-squares estimators or the scan estimators decrease as these elements increase. The second concept of efficiency is a collective index for a group of estimates and its optimal value is unity. The estimators are collectively less efficient as EF(II) increases.

Table 1 and Tables 2A-2D are arranged in two sections. Cells in the first three columns of these tables contain parametric estimates, based on the three different approaches, and "Student's t-statistics" for linear hypotheses that individual parameters are zero. Only the t-statistics for Aitken's GLS estimators[13] are based on appropriate formulae within the context of our model. The "t-statistics" for the least-squares and scan estimators are based on the inappropriate formulae that would be used if the Yule-Slutsky effect of aggregation were ignored. It is convenient to describe these as quasi-t-statistics. Clearly, these are the values that are generally

Table 1. (PERRY) Wage Equations for U. S. Manufacturing Industry (1948-1969)

	OLS	GLS	SCAN $\rho = 0.68$	$EF(I)_0$	$EF(I)_\rho$
Const.	- .00322 (-.571)	-.012 (-.983)	.00623 (1.76)	1.17	1.18
$\dot{C}$	1.93 (11.13)	1.52 (5.19)	1.61 (5.92)	1.60	1.32
R	0.198 (3.62)	0.341 (3.00)	0.353 (3.23)	1.28	1.33
1/U	0.816 (5.68)	0.543 (2.22)	0.0794 (2.99)	1.82	1.75
ΔR	0.535 (3.48)	0.176 (0.94)	0.716 (3.91)	2.42	1.52
DK	-.0134 (-2.65)	-.00618 (-.065)	-.0105 (-1.27)	1.39	1.26
DG	-.00981 (-4.61)	.000111 (.037)	-.00747 (-1.60)	2.95	3.00
F(6,81)		120.83	125.53		

$EF(II)_0 = 25.34$; $EF(II)_\rho = 16.64$.

VARIABLE DEFINITION FOR THE PERRY MODEL

Dependent variable: annual percentage change in straight-time hourly earnings of production workers for total, durable and non-durable manufacturing, $\left[(W_t - W_{t-4}) / W_{t-4}\right]$.

$\dot{C}$: four quarter moving average of one quarter percentage change in the consumer price index $\left(\sum_{i=0}^{3} \frac{C_{t-i} - C_{t-1-i}}{C_{t-1-i}}\right)$, lagged one quarter.

1/U : reciprocal of the four quarter moving average of the unemployment rate (scaled by 100).

R : four quarter moving average of the annual profit rate (ratio of corporate earnings after taxes to stockholders equity), lagged one quarter, for total, durable and non-durable manufacturing.

ΔR : first difference of the profit rate series.

DK : dummy variable for Korean War.

DG : dummy variable for guideposts.

Table 2A. (BODKIN et al.) Wage Equations for Canadian Manufacturing Industry (1953 I - 1965 II)

	OLS	GLS	SCAN $\rho = 0.285$	$EF(I)_0$	$EF(I)_\rho$
Const.	-4.122	-6.403	-2.57	2.07	1.85
$\dot{P}$	0.377 (4.88)	0.250 (1.09)	0.421 (4.15)	1.39	1.29
$1/U^2$	10.427 (1.55)	-5.514 (-0.39)	14.0 (1.65)	2.53	2.19
(Z/Q)	0.053 (2.95)	0.077 (1.93)	0.05 (2.26)	2.08	1.85
$\dot{W}_{us}$	0.432 (3.95)	0.444 (1.46)	0.399 (2.91)	1.47	1.32
$\dot{W}_{t-4}$	-0.092 (-2.28)	0.009 (0.12)	-0.179 (-2.42)	2.20	4.41
F(5,44)	44.335	5.419	41.89		

$EF(II)_0 = 10.75$; $EF(II)_\rho = 19.98$.

Table 2B. (BODKIN et al.)

	OLS	GLS	SCAN $\rho = 0.47$	$EF(I)_0$	$EF(I)_\rho$
Const.	-4.566	-7.226	-2.05	2.11	1.66
$\dot{P}$	0.458 (5.38)	0.276 (1.19)	0.510 (3.77)	1.36	1.21
$1/U^2$	22.839 (3.34)	4.730 (0.38)	25.4 (2.73)	2.56	1.89
(Z/Q)	0.066 (3.29)	0.096 (2.50)	0.0635 (2.36)	2.22	1.84
$\dot{W}_{t-4}$	-0.124 (-2.74)	-0.021 (-0.26)	-0.270 (-2.67)	2.15	4.69
F(4,45)	38.907	6.092	46.35		

$EF(II)_0 = 7.43$; $EF(II)_\rho = 14.36$.

Table 2C. (BODKIN et al.)

	OLS	GLS	SCAN $\rho = 0.35$	$EF(I)_0$	$EF(I)_\rho$
Const.	0.807	0.714	0.762	1.28	1.37
$\dot{P}$	0.428 (5.26)	0.238 (1.00)	0.483 (4.33)	1.33	1.24
$1/U^2$	22.931 (4.07)	4.079 (0.30)	26.2 (3.52)	1.88	1.62
$\dot{W}_{us}$	0.495 (4.26)	0.633 (2.13)	0.451 (3.03)	1.57	1.34
$\dot{W}_{t-4}$	-0.100 (-2.31)	0.064 (0.82)	-0.208 (-2.47)	2.53	5.11
$F(4,45)$	45.480	5.516	48.21		

$EF(II)_0 = 5.21$; $EF(II)_\rho = 10.82$

Table 2D. (BODKIN et al.)

	OLS	GLS	SCAN $\rho = 0.51$	$EF(I)_0$	$EF(I)_\rho$
Const.	1.809	2.169	1.13	1.23	1.64
$\dot{P}$	0.541 (5.99)	0.274 (1.12)	0.592 (4.07)	1.28	1.18
$1/U^2$	41.590 (10.02)	24.114 (2.35)	42.8 (7.06)	1.43	1.27
$\dot{W}_{t-4}$	-0.141 (-2.85)	0.038 (0.47)	-0.301 (-2.71)	2.35	5.23
F(3,46)	39.760	5.423	55.98		

$EF(II)_0 = 3.38$; $EF(II)_\rho = 8.80$

Table 2E. (BODKIN et al.) Efficiency $EF(I)_\rho$ for ρ equal to the alternative values 0.375 and 0.75.

ρ	0.375	0.750	0.375	0.750	0.375	0.750	0.375	0.750
Const.	1.75	1.45	1.77	1.47	1.36	1.40	1.57	1.87
$\dot{P}$	1.26	1.64	1.25	1.66	1.23	1.65	1.20	1.66
$1/U^2$	2.01	1.97	2.07	2.07	1.60	1.72	1.34	1.67
(Z/Q)	1.75	1.50	1.97	1.66	----	----	----	----
$\dot{W}_{us}$	1.25	1.20	----	----	1.32	1.25	----	----
$\dot{W}_{t-4}$	4.47	5.60	4.62	5.96	5.12	6.17	5.10	6.37

VARIABLE DEFINITION FOR THE BODKIN *et al.* STUDY

Dependent variable: annual percentage change in average hourly earnings of production workers in Canadian manufacturing industry, $\left[(W_t - W_{t-4}) \;/\; W_{t-4}\right]$.

$\dot{P}$: four quarter moving average of annual percentage change in consumer price index, $\left[1/4 \sum_{i=0}^{3} \left(\frac{P_{t-i} - P_{t-4-i}}{P_{t-4-i}}\right)\right]$.

$1/U^2$: squared reciprocal of a four quarter moving average of a two quarter average of the Canadian unemployment rate, $\left[1/8\ U_t + 1/4 \sum_{i=0}^{3} U_{t-i} + 1/8\ U_{t-4}\right]^{-2}$

(Z/Q) : four quarter moving average of the profit markup on output (index of corporate profits before tax divided by manufacturing production index), lagged two quarters.

$\dot{W}_{us}$: four quarter moving average of the annual percentage change in average hourly earnings in U. S. manufacturing expressed in U. S. dollars, $\left(1/4 \sum_{i=0}^{3} \frac{W_{us_{t-i}} \quad W_{us_{t-4-i}}}{W_{us_{t-4-i}}}\right)$.

$\dot{W}_{t-4}$: dependent variable lagged four quarters.

published. The final two columns contain our first index for the efficiencies of least-squares and scan estimators as compared with the relatively-efficient GLS estimators. Lower parts of the tables contain values for the second index (the generalized variances) for these efficiencies and calculated "F-statistics" for the linear hypothesis that all parameters are zero apart from the constant. Again the prefix quasi should be added to the entries for the least-squares and scan values for these statistics. Table 1 illustrates the Perry model whereas Tables 2A-2D illustrate four different specifications of Bodkin et al. In each case, the value of ρ in the autoregressive transformation is chosen by the Hildreth-Lu scan procedure. Table 2E contains values for the first index of efficiency (the pairwise comparisons of variances for individual parameters) when the choice of ρ is restricted to the two specific alternatives considered by Bodkin et al. Their selection of 0.375 appears correct within the framework of this simple choice, although not all of the evidence favours their decision.

Although interpretation of these tabulations appears straightforward, certain points are worthy of emphasis. First, inferences with respect to the significance of particular parameters are substantially inappropriate, if they are derived from least-squares or scan quasi-t-statistics, for both models although Perry's model does appear to be more robust in the presence of this autocorrelated mis-specification. Use of the scan procedure for Perry's model does revise two incorrect inferences based on the least-squares alternative; namely, the insignificance of the two dummy variables for guide-posts and the period of the Korean conflict. Second, application of the auto-regressive transformations do not necessarily improve efficiency of estimates even if the problem of bias is ignored. For example, two of Perry's six variables have least-squares estimates that are more efficient than their scan alternatives even though the least-squares estimates are collectively less efficient than these alternatives. In all of the specifications of Bodkin et al., the principal source of inefficiency stems from the presence of the lagged dependent variable. The inefficiency of scan estimates for this parameter is sufficiently large to make least-squares estimators collectively more efficient. Third, the losses in efficiency arising from a failure to use the GLS procedure are substantial in both models whether individual estimates or collective performance are considered.

Conclusions

The structure of the labour market suggests that the errors for models of Phillips' curves are characterized by moving-average processes, rather than the simple first-order autoregressive process which is the basis for the popular practice of autoregressive transformations. In the simplest model of annual increments, it is questionable whether the scan procedure is usually better than the least-squares method. The scan procedure is known to be inferior in the context of the specifications of Bodkin et al. since it introduces asymptotic bias. In any case, it is not clear why the slight addition in computational costs associated with use of the GLS procedure should prevent its widespread adoption. These costs are trivial when weighed against those attributable to invalid inferences and losses in efficiency that stem from the use of either the least-squares method or any of the techniques involving autoregressive transformations. Finally, the variability in empirical results suggests strongly that we should always attempt to justify the adoption of particular assumptions with respect to errors by reference to the source of the data and the institutional framework by which they are generated.

Footnotes

1. A definitive statement in this area is provided by Rao and Mitra, [1971, Ch. 8].

2. Bodkin et al., p. 154.

3. Bodkin et al., P. 125, footnote 1.

4. Perry [1966], pp. 30-31, provides the clearest exposition of this model in the economic literature. Many authors take the particular specifications as given and do not attempt to either justify or explain them. Kuh [1967] and Sargan [1964] provide important exceptions. These two studies have strongly influenced our thinking. See especially Sargan, p. 36.

5. This assumption is relaxed when the presence of lagged dependent variables is discussed in the context of the approach of Bodkin et al.

6. An appropriate statistical framework is provided by Rowley and Wilton [1971b, 1971c]. Some criticisms of fixed weights are valid. They are to be found in Rowley and Wilton [1971a].

7. Linearity is defined in terms of the vector of observations for changes in the aggregative wage-index Gy.

8. Perry [1966], p. 31. This quotation has been modified so that its language is consistent with our exposition but its meaning has been strictly preserved.

9. We return to the question of contemporaneous lack of correlation later.

10. Perry [1966], p. 52.

11. Bodkin et al., p. 160.

12. The problems of lagged dependent variables and simultaneity are ignored throughout this section. This myopia permits efficiencies to be discussed in terms of variances rather than mean-squared-errors. Qualifications to the tabulated results will probably suggest that they understate the losses in efficiencies of alternative estimators.

13. The problem of stochastic regressors must again be pointed out for the specifications of Bodkin et al.

REFERENCES

BODKIN, R.G., et al., *Price Stability and High Employment: The Options for Canadian Economic Policy* (Ottawa: Queen's Printer, 1966).

COCHRANE, D., and G. H. ORCUTT, "Application of least Squares Regression to Relationships Containing Autocorrelated Error Terms", *Journal of the American Statistical Association*, Vol. 46, 1949, pp. 32-61.

DICKS-MIREAUX, L.A., and J.C.R. DOW, "The Determinants of Wage Inflation: United Kingdom, 1946-56", *Journal of the Royal Statistical Society*, Series A, Vol. 22, Part II, 1959, pp. 145-184.

GOLDBERGER, A.C., *Econometric Theory* (New York: John Wiley, 1964).

HILDRETH, C., and J. LU, *Demand Relations with Autocorrelated Disturbances*, Michigan State University Agricultural Experimental Station, Technical Bulletin No. 276, 1960.

JOHNSTON, J., *Econometric Methods* (New York: McGraw-Hill, 1971), Second Edition.

KADIYALA, K.R., "A Transformation Used to Circumvent the Problem of Autocorrelation", *Econometrica*, Vol. 36, No. 1, January 1968, pp. 93-96.

KUH, E., "A Productivity Theory of Wage Levels - An Alternative to the Phillips Curve", *Review of Economic Studies*, 1967, Vol. 34, pp. 333-360.

PERRY, G.L., *Unemployment, Money Wage Rates and Inflation* (Cambridge: The M.I.T. Press, 1966).

__________, "Inflation and Unemployment", reprint from *Savings and Residential Financing: 1970 Conference Proceedings*, (Washington: The Brookings Institution, 1970).

RAO, C.R., and S.K. Mitra, *Generalized Inverse of Matrices and its Applications* (New York: Wiley, 1971).

ROWLEY, J.C.R., and D.A. WILTON, "Wage Determination: The Use of Instrumental Assumptions" (1971a). forthcoming in *International Economic Review*.

__________, "Quarterly Models of Wage Determination: Some New Efficient Estimates" (1971b), forthcoming in *American Economic Review*.

__________, "Moving-Average Transformations in Classical Linear Models" (1971c), forthcoming in *Metroeconomica*.

SARGAN, J.D., "Wages and Prices in the United Kingdom: A Study in Econometric Methodology", *Econometric Analysis for National Economic Planning*, (eds.) P. E. Hart, G. Mills, and J.K. Whitaker (London: Butterworths, 1964).